

1 SYSTEM AND METHOD FOR LANGUAGE VARIATION GUIDED
2 OPERATOR SELECTION

3
4 CROSS-REFERENCE TO RELATED CO-PENDING APPLICATION

5 This application may relate to co-pending U.S. Patent Application Serial No.
6 10/676995, entitled "System And Method For Operator Assisted Automated Call
7 Handling," filed on September 30, 2003, by Lin, and co-pending U.S. Patent
8 Application PDNo. 200310012, entitled "System And Method For Extracting
9 Demographic Information," filed on January 30, 2004, by Yacoub et al. These related
10 applications are commonly assigned to Hewlett-Packard of Palo Alto, CA.

11
12 BACKGROUND OF THE INVENTION

13 1. Field of the Invention

14 The present invention relates generally to systems and methods for call
15 handling, and more particularly to language variation guided operator selection.

16 2. Discussion of Background Art

17 Automated call handling systems, such as Interactive Voice Response (IVR)
18 systems, using Automatic Speech Recognition (ASR) and Text-to-speech (TTS)
19 software are increasingly important tools for providing information and services in a
20 more cost efficient manner. IVR systems are typically hosted by a server that includes
21 an array of Digital Signal Processors (DSPs), and enable users to interact with
22 corporate databases and services over a telephone using a combination of voice
23 utterances and telephone button presses. IVR systems are particularly cost effective
24 when a large number of users require data or services that are very similar in nature
25 and thus can be handled in an automated manner, often providing a substantial cost
26 savings due to a need for fewer human operators.

1 In an ideal situation, an IVR system would be able to automatically guide a
2 user through an entire transaction using only predefined dialogs, without any human
3 interference. In reality, speech recognition technology is still limited in some ways
4 and so users still need to be connected to human operators. Such operators, however,
5 can find understanding the user just as difficult. This may be especially true in
6 countries with a significant language, dialect, or accent diversity. Such user language
7 variations often present such a significant challenge to telephone operators, that it may
8 have an adverse effect on the call center's effectiveness and cost efficiency.

9 Call center effectiveness is affected if the user becomes frustrated at having to
10 repeat or rephrase often. Sometimes the user can feel insulted if believing that they
11 can fluently speak the language, the human operator still can not understand them.
12 This is an example of language-understanding mismatch that can have a negative
13 business impact.

14 Call center cost efficiency is affected if transactions take longer to finish due
15 to language difficulties thereby requiring more resources such as higher phone bills
16 and a greater number of voice operators that must be hired to keep call waiting times
17 below a predetermined level.

18 In response to the concerns discussed above, what is needed is a system and
19 method for operator selection that overcomes the problems of the prior art.
20

1 BRIEF DESCRIPTION OF THE DRAWINGS

2 Figure 1 is a dataflow diagram of one embodiment of a system for language
3 variation guided operator selection within an interactive voice response system;

4 Figure 2 is a root flowchart of one embodiment of a method for language
5 variation guided operator selection within an interactive voice response system; and

6 Figure 3 is a flowchart of one embodiment of a method for language variation
7 guided operator selection within an interactive voice response system.

8

1 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

2 The present invention provides a number of solutions for enabling human
3 operators to provide call center services to contacts having a wide diversity of
4 languages. Such language deviations are herein defined to include not only
5 completely different languages, but also variations in accent, dialect, and so on within
6 a single language or language family. Improved methods for analyzing a contact's
7 language variations and a method for monitoring operator performance with respect to
8 such language variations and managing call center resource typologies are also
9 presented.

10 Using the present invention, call center operators are assigned based on their
11 language and cultural strengths, reducing the chance of a contact-operator mismatch,
12 thereby permitting the dialog between the contact and operator to proceed more
13 smoothly and the transaction to be completed more quickly. Also, those calling in to
14 contact the present invention are likely to be happier since they will be better
15 understood. Culture barriers can also be reduced when contacts are routed to
16 operators who understand not only their language, but also the contact's idiomatic
17 usage, conversation style, and perhaps even sense of humor.

18 The present invention thus enhances contact satisfaction and call center
19 performance, especially in countries having a great diversity of languages and
20 language variations.

21
22 Figure 1 is a dataflow diagram of one embodiment of a call handling system
23 102 for language variation guided operator selection. The call handling system 102 of
24 the present invention preferably provides some type of voice interactive information
25 management service to a set of contacts. Anticipated information services include
26 those associated with customer response centers, enterprise help desks, business

1 generation and marketing functions, competitive intelligence methods, as well as
2 many others. Contacts may be customers, employees, or any party in need of the call
3 center's services.

4 To begin a contact 104 enters into a dialog with the call handling system 102.
5 While the dialog typically begins once a dialog manager 106 connects the contact 104
6 to an Interactive Voice Response (IVR) module 108 through a dialog router 110,
7 alternative dialogs could route the contact 104 directly to a human operator 112, 114,
8 or 116. The IVR module 108 provides an automated interface between the contact's
9 104 speech signals and the system's 102 overall functionality. To support such an
10 interface with the contact 104, the IVR module 108 may include a Text-To-Speech
11 (TTS) translator, Natural Language Processing (NLP) algorithms, Automated Speech
12 Recognition (ASR), and various other dialog interpretation (e.g. a Voice-XML
13 interpreter) tools.

14 As part of the dialog, the IVR module 108 receives information requests and
15 responses from the contact 104 which are then stored in a contact database 118.
16 While the IVR module 108 enables the system 102 to exchange information with the
17 contact 104 in a very efficient manner, from time to time the dialog can be interrupted
18 due to an inability of the IVR module 108 to correctly interpret the contact's 104
19 requests or responses.

20 In such cases, routing the contact 104 to an operator 112, 114, or 116 who can
21 correctly interpret the contact's 104 requests or responses is required in order to
22 successfully complete the dialog. Preparations for such a connection between the
23 contact 104 and an operator 112, 114, or 116 are preferably completed in parallel with
24 the ongoing dialog between the contact 104 and the IVR module 108 so that the
25 contact 104 can be connected to an appropriate operator as soon as needed. Such an
26 implementation of the present invention is described below. However, in alternate

embodiments, the preparations can be delayed until a connection between the contact 104 and an operator is actually requested.

Preparations for identifying a language variation spoken by the contact 104 begin when a set of predefined language variations are stored in a language database 120. Some ways of selecting an appropriate set of language variations include: 1) those that experience has shown are particularly challenging to understand; 2) those variations that are different enough so that they can be reliably distinguished; and 3) those languages, dialects, accents, and so on which the call center where the system 102 is located has typically encountered. For example, if the system 102 serves residents of the United States, the following language variations within spoken English may be encountered: Standard English, African American, Southern, Indian, Chinese, Japanese, Arabic, French, and so on.

The contact language classifier 122 generates a set of confidence scores indicating a probability that the contact's speech signals within the dialog fall within any of the predefined language variations. The language classifier 122 generates such confidence scores based on prior training on the set of predefined language variations using either neural net, clustering (based on various voice parameters, including cepstral coefficients, and other voice signal processing elements), or other training methods. The contact's speech signals analyzed by the language classifier 122 can either be captured in real time from the contact 104 during the dialog, or retrieved from the contact database 118.

While the language classifier 122 can assign a confidence score ("p") to each language variation based only on various speech features extracted from the speech

1 signal, preferably contact language variation classification is further improved using
2 an inverse-distance weighting scheme.

3 The inverse-distance weighting scheme works by first selecting one of the
4 language variations as an origin. Next, a distance between the origin language
5 variation and each of the other language variations is calculated. Third, these
6 distances are normalized with respect to the origin language variation. Then each
7 normalized distance is multiplied by its respective confidence score. Finally, the
8 results are summed to yield an inverse-distance weighted confidence score for the
9 origin language variation.

10 Additional inverse-distance weighted confidence scores are then calculated by
11 letting each of the other language variations take a turn as the origin, until all of the
12 language variations have been selected as the origin.

13 Preferably the distance normalization factor is either adaptive or set by a large
14 corpus of ground truth. Also individual cities can be further sub-clustered based on:
15 regional languages, the happenstance of history (e.g., Miami may have many displaced
16 northerners, West Texas was settled much later than Louisiana), as well as other
17 language sub-clustering variations. International languages also can be viewed as
18 having a distance between them. For example, Korea and Japan are geographically
19 closer than Mexico and Estonia.

20
21 Those language variations having an inverse-distance weighted confidence
22 score above a predetermined value are associated with the contact 104 and stored in
23 the contact database 118. While preferably only the highest inverse-distance weighted
24 confidence score is associated with the contact 104, in certain difficult or borderline
25 cases two or more language variations may have similar inverse-distance weighted
26 confidence score for the contact 104 and thus may be stored in the contact database as

1 well. Subsequent phone calls by the contact 104 to the call handling system 102 may
2 help distinguish these similar confidence scores further or be used to generate a report
3 indicating that the set of predefined language variations could be improved upon to
4 better distinguish the contacts.

5 The distance used in the inverse-distance weighting scheme can either be a
6 physical distance or a virtual distance. Physical distance associates various
7 geographic locations having known language variations. The contact's 104
8 geographic location can be approximated either based on the phone number the
9 contact is calling from or an analysis of the contact's 104 language. Virtual distances
10 however are based on various voice signature parameters of the contact 104 as
11 measured using one or more of metrics, such as neural nets, voice pre-processing
12 signal analysis, and so on.

13 An example of an inverse-distance weighted confidence score calculation is
14 now discussed. In this example, the contact's 104 speech signal is compared with the
15 following predetermined language variations labeled as: Chicago, Milwaukee,
16 Minneapolis, and Saint Louis. First a normal set of confidence scores ("p") are
17 calculated with respect to the contact's 104 speech signal. Then, a set of distances
18 ("d") are calculated for each language variation with respect to each other language
19 variation. Finally the inverse-distance weighting scheme is applied using the
20 distances to improve the confidence scores.

21 The inverse-distance weighted confidence score, where Chicago is selected as
22 a language variation origin is calculated first. So if the confidence scores and
23 distances are: Chicago $p=0.45$; Milwaukee $p=0.4$, geographic distance from Chicago =
24 100 km, $1/d=0.01$; Minneapolis, $p=0.3$, geographic distance from Chicago = 400 km,
25 $1/d=0.0025$; Saint Louis, $p=0.25$, geographic distance from Chicago = 300 km,
26 $1/d=0.0033$; and the normalization factor is 0.01, which is the distance ratio from

1 Chicago to Milwaukee, then the inverse-distance weighted confidence score for
 2 Chicago is $= 0.45 + 0.4*(0.01/0.01) + 0.3*(0.0025/0.01) + 0.25*(0.0033/0.01) =$
 3 1.008.

4 The inverse-distance weighted confidence score, where Milwaukee is selected
 5 as a language variation origin is calculated next. If the confidence scores and
 6 distances are: Milwaukee $p=0.4$; Chicago $p=0.45$, geographic distance from
 7 Milwaukee = 100 km, $1/d=0.01$; Minneapolis, $p=0.3$, geographic distance from
 8 Milwaukee = 300 km, $1/d=0.0025$; Saint Louis, $p=0.25$, geographic distance from
 9 Milwaukee = 400 km, $1/d=0.0033$; and the normalization factor is 0.01, which is the
 10 distance ratio from Milwaukee to Chicago, then the inverse-distance weighted
 11 confidence score for Milwaukee is $= 0.40 + 0.45*(0.01/0.01) + 0.3*(0.0033/0.01) +$
 12 $0.25*(0.0025/0.01) = 1.0125$

13 The inverse-distance weighted confidence scores for Minneapolis and Saint
 14 Louis would then also be calculated in a same manner.

15 As can be seen from this example, inverse-distance weighted confidence score
 16 for Milwaukee turns out to be higher than the inverse-distance weighted confidence
 17 score Chicago, even though Chicago had a higher basic confidence score ("p"). This
 18 shows how use of the inverse-distance weighting scheme can improve contact
 19 language variation predictions.

20

21 The contact's 104 language variation is retrieved from the contact database
 22 118 each time the contact 104 calls the call handling system 102. Retrieval obviates
 23 the need to recalculate the contact's 104 inverse-distance weighted confidence score
 24 each time the contact 104 calls. To retrieve the correct inverse-distance weighted
 25 confidence score, the contact 104 preferably identifies themselves to the system 102
 26 by entering in an account number.

1
2 In preparing to assign a best operator 112, 114, or 116 to the contact's 104
3 call, each operators 112, 114, and 116 skill level or performance with respect to each
4 of the predetermined language variations is preferably determined using, an operator
5 language performance module 124. In the embodiment described below, the operator
6 performance module 124 calculates each operator's 112, 114, and 116 skill level for
7 each of the language variations in real time, as each operator 112, 114, or 116 dialogs
8 with a contact. However, those skilled in the art will recognize that other
9 embodiments may first rate the operators on a series of test contacts each speaking
10 with a different language variation. Such an embodiment may be preferred if the
11 importance of not misinterpreting real contacts is quite high.

12 To begin, the operator performance module 124 assigns each operator an
13 initial skill level with respect to each of the language variations. After the number of
14 operators who performance has been rated with respect to a particular language
15 variation reaches a predetermined threshold (preferably 10), then an average skill level
16 for all of the rated operators is calculated and replaces the initial skill level assigned to
17 an operator new to the language variation.

18
19 Next, the operator performance module 124 rates the difficulty of each dialog
20 between a contact and an operator. There are several ways in which the difficulty
21 score for a dialog can be determined. One way is for an operator to self rate
22 themselves regarding how difficult a dialog was with a contact who speaks the
23 language variation was after the dialog is complete.

24 Another way defines a set of dialog key words indicating communication
25 difficulties, such as "sorry", "repeat", "say it again", "pardon" and "excuse me". Then
26 the operator is rated based on how many of the key words the operator spoke in a

1 dialog with a contact who speaks the language variation. The keywords can be
2 detected using automatic speech recognition software running in the background. The
3 number of such words or phrases reflects how difficult the dialog was.

4 A third way measures a time spent during the dialog between the contact and
5 the operator on each menu, form, stage, or portion of the dialog, together with a
6 number of words uttered during each portion. The operator is then rated based on the
7 time spent and number of words spoken. More difficult dialogs will take longer and
8 have a greater word count. Lastly, some combination of the above techniques can be
9 used as well. Weighted linear summation of scores obtained from these techniques
10 can be used to potentially increase the difficulty score's reliability.

11 The difficulty rating for the language variation of the dialog between the
12 contact and the operator is associated with the operator and stored in an operator
13 database 126. The operator performance module 124 determines each operator's
14 overall skill level with respect to each language variation by averaging all of the
15 operator's difficulty ratings within each language variation. Those skilled in the art
16 recognize other ways of rating operator performance.

17 After the above preparations have been completed, the call handling system
18 102 is now ready to match the contact 104 with one of the operators 112, 114, or 116
19 should such a connection be requested.

20
21 The dialog manager 106 receives a request from either the IVR module 108 or
22 the contact 104 to connect the contact 104 with an operator 112, 114, or 116. The
23 IVR module 108 initiates the request if the IVR module 108 detects that a same
24 question or response has been presented to or received from the contact 104 a
25 predetermined number of times. The contact 104 can initiate a request by pressing a
26 hot key (for example, “#” key) or speaking a special voice command, such as “Help”.

1 Next, the dialog manager 106 retrieves the language variation associated with
2 the contact 104 from the contact database 118, and identifies from the operator
3 database 126 those operators having a skill level above a predetermined level for the
4 contact's language variation.

5 The dialog manager 106 commands the dialog router 110 to transfer the
6 contact 104 to that operator that is available and has a highest skill level for the
7 contact's language variation. Alternatively, the dialog router 110 could queue the
8 contact 104 with a soon to be available operator having a highest skill level for the
9 contact's language variation. If more than one operator has equally high skill levels,
10 then an arbitrary selection is made as to which operator to connect the contact 104
11 with.

12 Other information known about the contact 104 and the operators 112, 114,
13 and 116 can be used to help identify which operator would be best suited to handle the
14 contact's 104 call. Such other information may include: an operator's second
15 language, and cultural background. For example, if the contact 104 speaks English
16 with German-accent, then the contact 104 is preferably routed to an operator who both
17 skillfully understands both German-accented English and German, and who knows the
18 German culture. Such a cultural understanding on the part of the operator can yield a
19 deeper understanding of the contact 104 than would otherwise be possible even if the
20 operator knew the contact's language. For instance, interacting with a contact who is
21 originally from Germany may be different than that of a contact who is originally from
22 Italy.

23
24 The present invention's ability to access the language variations of contacts
25 and operator performance with respect to such variations, permits the present

invention to generate reports on how efficiently contacts are processed by the call handling system 102 and how the system's 102 typology might be modified.

The dialog manager 106 retrieves data from the contact database 118 and generates a report on which language variations the contacts calling the call center aggregate into. For example, 30% of the contacts may have a southern accent, 40% may have a western accent, 20% may have an eastern accent, and 10% may have a mid-western accent.

The dialog manager 106 retrieves data from the operator database 126 and generates a report on the skill levels of the operators 112, 114, and 116 with respect to each of the language variations. For example, 10% of the operators may be reasonably skilled in southern accents, 60% may be skilled in western accents, 20% may be skilled in eastern accents, and 10% may be skilled in mid-western accents.

The dialog manager 106 generates a report on any language variations disparities between contacts calling the call center and the skill level of operators at the call center. Such disparities enable a call center manager to make decision on which operators to hire at various call centers, how call center resources might be better allocated, and whether contact calls should be routed to a different call center having a greater number of operators skilled in that contact's language variation.

Figure 2 is a flowchart of a one embodiment of a root method 200 for language variation guided operator selection. The method 200 begins in step 202 where a dialog between a contact and a call handling system is initiated. Next in step 204, a language variation spoken by the contact is identified. In step 206, a skill level with respect to the language variation for each operator within a set of operators is determined. In step 208, an operator whose skill level in the language variation is above a predetermined value is selected. Then in step 210 the dialog with the contact

1 is transferred to the operator. The root method 200 is discussed in further detail with
2 respect to Figures 3A and 3B below.

3
4 Figures 3A and 3B are a flowchart of one expanded embodiment 300 of the
5 root method 200 for language variation guided operator selection. In step 302, a
6 contact 104 enters into a dialog with the call handling system 102. In step 304, as part
7 of the dialog, the IVR module 108 receives information requests and responses from
8 the contact 104 that are then stored in a contact database 118.

9 Preparations for identifying a language variation spoken by the contact 104
10 begin in step 306, where a set of predefined language variations are stored in a
11 language database 120. In step 308, the contact language classifier 122 generates a set
12 of confidence scores indicating a probability that the contact's speech signals within
13 the dialog fall within any of the predefined language variations. Contact language
14 variation classification is further improved using an inverse-distance weighting
15 scheme, in step 310. In step 312, those language variations having an inverse-distance
16 weighted confidence score above a predetermined value are associated with the
17 contact 104 and stored in the contact database 118. In step 314, the contact's 104
18 language variation is retrieved from the contact database 118 each time the contact
19 104 calls the call handling system 102.

20 In step 316, the operator performance module 124 assigns each operator an
21 initial skill level with respect to each of the language variations. Next, in step 318, the
22 operator performance module 124 rates the difficulty of each dialog between a contact
23 and an operator. In step 320, the difficulty rating for the language variation of the
24 dialog between the contact and the operator is associated with the operator and stored
25 in an operator database 126. In step 322, the operator performance module 124
26 determines each operator's overall skill level with respect to each language variation

1 by averaging all of the operator's difficulty ratings within each language variation.

2 Those skilled in the art recognize other ways of rating operator performance.

3 After the above preparations have been completed, the call handling system
4 102 is now ready to match the contact 104 with one of the operators 112, 114, or 116
5 should such a connection be requested. So in step 324, the dialog manager 106
6 receives a request from either the IVR module 108 or the contact 104 to connect the
7 contact 104 with an operator 112, 114, or 116. Next, in step 326, the dialog manager
8 106 retrieves the language variation associated with the contact 104 from the contact
9 database 118, and identifies from the operator database 126 those operators having a
10 skill level above a predetermined level for the contact's language variation. In step
11 328, the dialog manager 106 commands the dialog router 110 to transfer the contact
12 104 to that operator that is available and has a highest skill level for the contact's
13 language variation.

14 In step 330, the dialog manager 106 retrieves data from the contact database
15 118 and generates a report on language variations into which the contacts calling the
16 call center aggregate. In step 332, the dialog manager 106 retrieves data from the
17 operator database 126 and generates a report on the skill levels of the operators 112,
18 114, and 116 with respect to each of the language variations. In step 334, the dialog
19 manager 106 generates a report on any language variations disparities between
20 contacts calling the call center and the skill level of operators at the call center.

21

22 While one or more embodiments of the present invention have been described,
23 those skilled in the art will recognize that various modifications may be made.

24 Variations upon and modifications to these embodiments are provided by the present
25 invention, which is limited only by the following claims.